UNIVERSIDADE ABERTA
www.uab.pt

Ministério da Educação e Ciência

U.C. 21044

Estatística para as Ciências Sociais

17 de junho de 2013

- INSTRUÇÕES -

- A prova é composta por **3** grupos de questões e respectivas alíneas, contém 4 página(s) e termina com a palavra **FIM**.
- Verifique o seu exemplar e, caso encontre alguma anomalia, dirija-se ao professor vigilante nos primeiros 15 minutos da prova, pois qualquer reclamação sobre defeito(s) de formatação e/ou de impressão não será aceite depois deste período.
- Todas as questões deverão ser respondidas no Caderno de Prova. Todos os cabeçalhos e espaços reservados à identificação, deverão ser preenchidos com letra legível. Utilize unicamente tinta azul ou preta.
- É permitido o uso de máquina de calcular. É fornecido um formulário na segunda página deste enunciado.
- O aluno pode arbitrar um valor adequado da resposta a uma alínea que não tenha respondido, caso este seja necessário para a resolução de uma alínea posterior.
- Verifique no momento da entrega da(s) folha(s) de ponto se todas as páginas estão rubricadas pelo vigilante. Caso necessite de mais do que uma folha de ponto, deverá numerá-las no canto superior direito.
- Não serão aceites folhas de ponto dobradas ou danificadas. Exclui-se, para efeitos de classificação, toda e qualquer resposta apresentada em folhas de rascunho.
- Se o enunciado não for utilizado para resposta, o aluno pode ficar com o exemplar.
- Tenha em atenção que a prova tem a duração máxima de **2 horas e 30 minutos**.

Cotação e critérios de correção:

- Clareza, correção, completude e justificação das respostas;
- Apresentação de fórmulas, tabelas e cálculos auxiliares;

COTAÇÕES																	
1. (6 valores)					2. (8.5 valores)							3. (5.5 valores)					
a)	b)	c)	d)	e)	a)	b)	c)	d)	e)	f)	g)	a)	b)	c)	d)	e)
1	1	1	1.5	1.5	1.5	0.5	2	2	0.5	1	1	1	1	1	1.5	1

FORMULÁRIO

ESTATÍSTICA DESCRITIVA

População de dimensão N

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i = \frac{X_1 + X_2 + \cdots + X_N}{N}; \qquad s = \sigma = \sqrt{s^{*2}};$$

$$s^{*2} = \sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(X_i - \bar{X})^2 = \frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \cdots + (X_N - \bar{X})^2}{N};$$

Amostra de dimensão n

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{x_1 + x_2 + \cdots + x_n}{n}; \qquad s = \sqrt{s^2};$$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2}{n-1};$$

Amostra de dim. n com dados agrupados em k classes de centro c_i com frequência n_i

$$\bar{x} \approx \frac{1}{n}\sum_{i=1}^{k} n_i c_i = \frac{n_1 c_1 + n_2 c_2 + \cdots + n_k c_k}{n}; \qquad s = \sqrt{s^2};$$

$$s^2 \approx \frac{1}{n-1}\sum_{i=1}^{k} n_i (c_i - \bar{x}) = \frac{n_1(c_1 - \bar{x})^2 + n_2(c_2 - \bar{x})^2 + \cdots + n_k(c_k - \bar{x})^2}{n-1};$$

$$\text{Mediana} = \begin{cases} x_{\left(\frac{n+1}{2}\right)}, & n \text{ é ímpar;} \\ \dfrac{x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}}{2}, & n \text{ é par;} \end{cases} \qquad \begin{aligned} Q_1 &= \text{ Mediana da } 1^a \text{ sub-amostra;} \\ Q_3 &= \text{ Mediana da } 2^a \text{ sub-amostra;} \end{aligned}$$

$$\text{Amplitude} = \max(x_i) - \min(x_i); \qquad \text{AIQ} = Q_3 - Q_1;$$

REGRESSÃO E CORRELAÇÃO

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{(n-1)s_x s_y} = \frac{(x_1 - \bar{x})(y_1 - \bar{y}) + (x_2 - \bar{x})(y_2 - \bar{y}) + \cdots + (x_n - \bar{x})(y_n - \bar{y})}{(n-1)s_x s_y}$$

Regressão linear: $\hat{y} = a + bx$; $\quad r^2 \equiv$ percentagem de explicação da variação de y por x;

PROBABILIDADES

$$P(A) = \frac{\text{no. de casos de A}}{\text{no. de casos possíveis}}; \qquad P(\bar{A}) = 1 - P(A); \quad P(A|B) = \frac{P(A \cap B)}{P(B)};$$

$$A \subset B \Rightarrow P(A) \leq P(B); \qquad P(A - B) = P(A) - P(A \cap B);$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B);$$

$$A \text{ e } B \text{ independentes } \Leftrightarrow P(A|B) = P(A) \Leftrightarrow P(A \cap B) = P(A) \cdot P(B);$$

1. Foram recolhidos o número de descendentes e o respectivo ordenado líquido mensal (em euros) de uma amostra de agregados familiares, consoante a tabela seguinte:

Nº descendentes (x_i)	0	1	0	2	5	2	1	1
Ordenado (y_i)	1430	1860	2520	1700	3300	2900	1850	2890

a) Indique um contexto possível para a recolha desta amostra, indicando claramente o objetivo de estudo e o tipo de recolha realizado.

b) Nesse contexto identifique a amostra e a população.

c) Classifique as variáveis número de descendentes e ordenado quanto ao tipo de dados e à continuidade.

d) Determine a tabela de frequências absolutas e relativas (simples e acumuladas) da variável nº de descendentes.

e) Esboce num mesmo gráfico, os gráficos de barras de frequências relativas simples e acumuladas para a variável número de descendentes.

2. Considere os mesmos dados da tabela do exercício 1.

a) Calcule as medidas de tendência central possíveis para a variável número de descendentes.

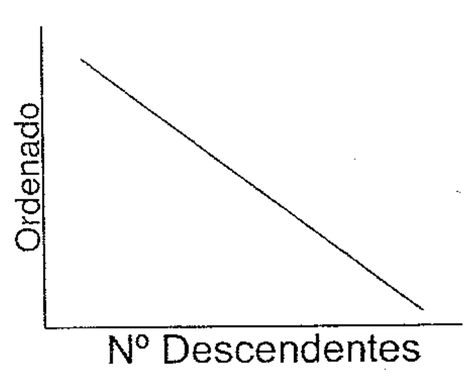
b) O que pode dizer sobre a simetria da distribuição da variável número de descendentes.

c) Calcule o desvio padrão da variável número de descendentes.

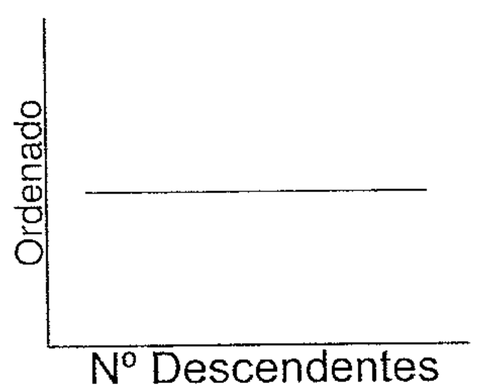
d) Calcule a correlação entre a variável número de descendentes e a variável ordenado, sabendo que a variância amostral da variável ordenado é $s_y^2 = 467026.8$.

e) Que conclui da alínea anterior?

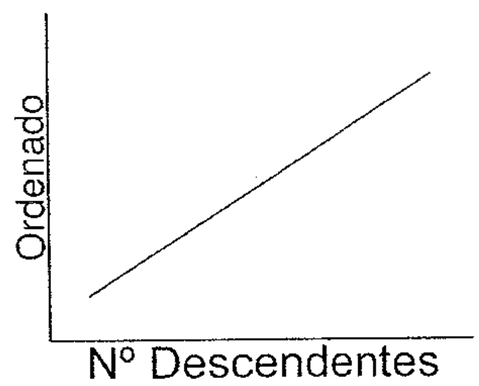
f) Indique, justificando, qual das figuras seguintes melhor representa a reta de regressão linear entre o número de descendentes e o ordenado.



i)



ii)



iii)

g) Sabendo que a reta de regressão entre as duas variáveis é dada por

$$\hat{y} = 1920.8 + 256.94x,$$

calcule a estimativa para o ordenado médio de um agregado com 3 descendentes.

3. Numa empresa sabe-se que 30% dos funcionários são fumadores e que 75% têm filhos. Parte-se do príncipio que as variáveis são independentes.

a) Represente graficamente de forma que considere adequada as frequências relativas desta população em relação à variável 'fumador'.

b) Determine a percentagem de funcionários que simultaneamente são fumadores e têm filhos.

c) De entre os fumadores, qual a percentagem de funcionários que têm filhos.

d) Esboce o diagrama de Venn que caracteriza os acontecimentos 'fumador' e 'ter filhos', indicando a legenda da figura.

e) Qual a percentagem de pessoas que simultaneamente não fumam e não têm filhos.

FIM